

# What Boolean Algebras Can Teach Us About Weighted Model Counting

Paulius Dilkas

3rd June 2020

## 1 Introduction

### Contributions.

- WMC defines a measure over a BA.
- WMC with weights on literals imposes an independence assumption. (Measures are ‘slightly’ more expressive than WMC with weights on models because they apply to non-atomic BAs.)
- A BA can be augmented with new literals in order to support any measure.
- (Maybe) a lower bound on the number of new literals needed in order to support any measure.
- Alternatively, one can use coproducts and pushouts to define a BA with precisely the right independence and conditional independence conditions. (This requires a relaxed version of WMC.)
- This results in a smaller problem for WMC algorithms (w.r.t. both the number of literals and the length of the theory) and is optimal for, e.g., Bayesian networks.
- (Maybe) this results in faster inference (?)

### Notable previous/related work.

- Hailperin’s approach to probability logic [9]
- Nilsson’s (somewhat successful) probabilistic logic [14]
- Logical induction: a big paper with a good overview of previous attempts to assign probabilities to logical sentences in a sensible way [7]
- Measures on Boolean algebras
  - On possibility and probability measures in finite Boolean algebras [2]
  - Representation of conditional probability measures [12]

## 2 Preliminaries

**Definition 1.** A *Boolean algebra* (BA) is a tuple  $(\mathbf{B}, \wedge, \vee, \neg, 0, 1)$  consisting of a set  $\mathbf{B}$  with binary operations *meet*  $\wedge$  and *join*  $\vee$ , unary operation  $\neg$  and elements  $0, 1 \in \mathbf{B}$  such that the following axioms hold for all  $a, b, c \in \mathbf{B}$ :

- both  $\wedge$  and  $\vee$  are associative and commutative;

- $a \vee (a \wedge b) = a$ , and  $a \wedge (a \vee b) = a$ ;
- 0 is the identity of  $\vee$ , and 1 is the identity of  $\wedge$ ;
- $\vee$  distributes over  $\wedge$  and vice versa;
- $a \vee \neg a = 1$ , and  $a \wedge \neg a = 0$ .

For clarity and succinctness, we will occasionally use three other operations that can be defined using the original three<sup>1</sup>:

$$\begin{aligned} a \rightarrow b &= \neg a \vee b, \\ a \leftrightarrow b &= (a \wedge b) \vee (\neg a \wedge \neg b), \\ a + b &= (a \wedge \neg b) \vee (\neg a \wedge b). \end{aligned}$$

We can also define a partial order  $\leq$  on  $\mathbf{B}$  as  $a \leq b$  if  $a = b \wedge a$  (or, equivalently,  $a \vee b = b$ ) for all  $a, b \in \mathbf{B}$ . Furthermore, let  $a < b$  denote  $a \leq b$  and  $a \neq b$ . For the rest of this paper, let  $\mathbf{B}$  refer to the BA  $(\mathbf{B}, \wedge, \vee, \neg, 0, 1)$ . For any  $S \subseteq \mathbf{B}$ , we write  $\bigvee S$  for  $\bigvee_{x \in S} x$  and call it the *supremum* of  $S$ . Similarly,  $\bigwedge S = \bigwedge_{x \in S} x$  is the *infimum*. By convention,  $\bigwedge \emptyset = 1$  and  $\bigvee \emptyset = 0$ . For any  $a, b \in \mathbf{B}$ , we say that  $a$  and  $b$  are *disjoint* if  $a \wedge b = 0$ .

**Definition 2** ([10, 13]). An element  $a \neq 0$  of  $\mathbf{B}$  is an *atom* if, for all  $x \in \mathbf{B}$ , either  $x \wedge a = a$  or  $x \wedge a = 0$ . Equivalently,  $a \neq 0$  is an atom if there is no  $x \in \mathbf{B}$  such that  $0 < x < a$ . We say that  $\mathbf{B}$  is *atomic* if for every  $a \in \mathbf{B} \setminus \{0\}$ , there is an atom  $x$  such that  $x \leq a$ .

**Lemma 3** ([6]). For any two distinct atoms  $a, b \in \mathbf{B}$ ,  $a \wedge b = 0$ .

**Lemma 4** ([8]). The following are equivalent:

- $\mathbf{B}$  is atomic.
- For any  $x \in \mathbf{B}$ ,  $x = \bigvee_{\text{atoms } a \leq x} a$ .
- 1 is the supremum of all atoms.

**Lemma 5** ([8]). All finite BAs are atomic.

**Definition 6** ([5, 10]). A *measure* on  $\mathbf{B}$  is a function  $m: \mathbf{B} \rightarrow \mathbb{R}_{\geq 0}$  such that:

- $m(0) = 0$ ;
- $m(a \vee b) = m(a) + m(b)$  for all  $a, b \in \mathbf{B}$  whenever  $a \wedge b = 0$ .

If  $m(1) = 1$ , we call  $m$  a *probability measure*. Also, if  $m(x) > 0$  for all  $x \neq 0$ , then  $m$  is *strictly positive*.

**Definition 7** ([8]). Let  $\mathbf{A}$  and  $\mathbf{B}$  be BAs. A (*Boolean*) *homomorphism* from  $\mathbf{A}$  to  $\mathbf{B}$  is a map  $f: \mathbf{A} \rightarrow \mathbf{B}$  such that:

- $f(x \wedge y) = f(x) \wedge f(y)$ ,
- $f(x \vee y) = f(x) \vee f(y)$ ,
- $f(\neg x) = \neg f(x)$

for all  $x, y \in \mathbf{A}$ .

**Lemma 8** (Homomorphisms preserve order [8]). Let  $f: \mathbf{A} \rightarrow \mathbf{B}$  be a homomorphism between two BAs  $\mathbf{A}$  and  $\mathbf{B}$ . Then, for any  $x, y \in \mathbf{A}$ , if  $x \leq y$ , then  $f(x) \leq f(y)$ .

<sup>1</sup>We use  $+$  to denote symmetric difference because it is the additive operation of a Boolean ring.

**Lemma 9** ([16]). *For any  $a, b \in \mathbf{B}$ ,  $a \leq b$  if and only if  $a \wedge \neg b = 0$ .*

**Lemma 10** ([8]). *Let  $m: \mathbf{B} \rightarrow \mathbb{R}_{\geq 0}$  be a measure. Then for all  $a, b \in \mathbf{B}$ , if  $a \leq b$ , then  $m(a) \leq m(b)$ .*

**Definition 11** ([11]). Let  $S$  be a set, and let  $\mathbf{B}$  be a BA. Then  $\mathbf{B}$  is a *free BA over  $S$*  if there is a map  $S \rightarrow \mathbf{B}$  such that for any BA  $\mathbf{C}$  and map  $S \rightarrow \mathbf{C}$ , there is a unique homomorphism  $\mathbf{B} \rightarrow \mathbf{C}$  that makes

$$\begin{array}{ccc} S & \longrightarrow & \mathbf{B} \\ & \searrow & \vdots \\ & & \mathbf{C}. \end{array}$$

commute. A BA  $\mathbf{B}$  is *free* if  $S$  exists.

**Lemma 12** ([16]). *A finite BA is free if and only if it has  $2^{2^n}$  elements for some  $n \in \mathbb{N}$ . It then has  $2^n$  atoms and  $n$  generators.*

### 3 WMC as a Measure

**Definition 13.** Let  $\mathcal{L}$  be a propositional (or first-order) logic, and let  $\Delta$  be a theory in  $\mathcal{L}$ . We can define an equivalence relation on formulas in  $\mathcal{L}$  as

$$\alpha \sim \beta \quad \text{if and only if} \quad \Delta \vdash \alpha \leftrightarrow \beta$$

for all  $\alpha, \beta \in \mathcal{L}$ . Let  $[\alpha]$  denote the equivalence class of  $\alpha \in \mathcal{L}$  with respect to  $\sim$ . We can then let  $B(\Delta) = \{[\alpha] \mid \alpha \in \mathcal{L}\}$  and define the structure of a BA on  $B(\Delta)$  as

$$\begin{aligned} [\alpha] \vee [\beta] &= [\alpha \vee \beta], \\ [\alpha] \wedge [\beta] &= [\alpha \wedge \beta], \\ \neg[\alpha] &= [\neg\alpha], \\ 1 &= [\alpha \rightarrow \alpha], \\ 0 &= [\alpha \wedge \neg\alpha] \end{aligned}$$

for all  $\alpha, \beta \in \mathcal{L}$ . Then  $B(\Delta)$  is the *Lindenbaum-Tarski algebra* of  $\Delta$  [11, 17].

**Example 14.** Let  $\mathcal{L}$  be a propositional logic with  $p$  and  $q$  as its only atoms. Then  $L = \{p, q, \neg p, \neg q\}$  is its set of literals. Let  $w: L \rightarrow \mathbb{R}_{\geq 0}$  be the *weight function* defined by

$$\begin{aligned} w(p) &= 0.3, \\ w(\neg p) &= 0.7, \\ w(q) &= 0.2, \\ w(\neg q) &= 0.8. \end{aligned}$$

Let  $\Delta$  be a theory in  $\mathcal{L}$  with a sole axiom  $p$ . Then  $\Delta$  has two models, i.e.,  $\{p, q\}$  and  $\{p, \neg q\}$ . The *weighted model count* (WMC) [3] of  $\Delta$  is then

$$\sum_{\omega \models \Delta} \prod_{\omega \models l} w(l) = w(p)w(q) + w(p)w(\neg q) = 0.3.$$

The corresponding BA  $B(\Delta)$  can then be constructed using Definition 13. Alternatively, one can first construct the free BA generated by the set  $\{p, q\}$  and then take a quotient with respect to either the filter generated by  $p$  or the ideal<sup>2</sup> generated by  $\neg p$ .

Each element of  $B(\mathcal{L})$  can also be seen as a subset of the set of all models of  $\mathcal{L}$ , with 0 representing  $\emptyset$ , 1 representing the set of all (four) models, each atom representing a single model, and each edge going upward representing a subset relation. Thus, the Boolean-algebraic way of calculating the WMC of  $\Delta$  consists of:

<sup>2</sup>More details on these concepts can be found in many books on BAs [8, 11].

1. Identifying an element  $a \in B(\mathcal{L})$  that corresponds to  $\Delta$ .
2. Finding all atoms of  $B(\mathcal{L})$  that are ‘dominated’ by  $a$  according to the partial order.
3. Using  $w$  to calculate the weight of each such atom.
4. Adding the weights of these atoms.

This motivates the following definition of WMC generalised to BAs.

**Definition 15.** Let  $\mathbf{B}$  be an atomic BA, and let  $M \subset \mathbf{B}$  be its set of atoms. Let  $L \subset \mathbf{B}$  be such that every atom  $m \in M$  can be uniquely expressed as  $m = \bigwedge L'$  for some  $L' \subseteq L$ , and let  $w: L \rightarrow \mathbb{R}_{\geq 0}$  be arbitrary. The *weighted model count*  $\text{WMC}_w: \mathbf{B} \rightarrow \mathbb{R}_{\geq 0}$  is defined as

$$\text{WMC}_w(x) = \begin{cases} 0 & \text{if } x = 0 \\ \prod_{l \in L'} w(l) & \text{if } M \ni x = \bigwedge L' \\ \sum_{\text{atoms } a \leq x} \text{WMC}_w(a) & \text{otherwise} \end{cases}$$

for any  $x \in \mathbf{B}$ . Furthermore, we define the *normalised weighted model count*  $\text{NWMC}_w: \mathbf{B} \rightarrow [0, 1]$  as  $\text{NWMC}_w(x) = \frac{\text{WMC}_w(x)}{\text{WMC}_w(1)}$  for all  $x \in \mathbf{B}$ . For both  $\text{WMC}_w$  and  $\text{NWMC}_w$ , we will drop the subscript when doing so results in no potential confusion. Finally, we say that a measure  $m: \mathbf{B} \rightarrow \mathbb{R}_{\geq 0}$  is a *WMC measure* (or is *WMC-definable*) if there exists a subset  $L \subset \mathbf{B}$  and a weight function  $w: L \rightarrow \mathbb{R}_{\geq 0}$  such that  $m = \text{WMC}_w$ .

**Theorem 16.** WMC is a measure, and NWMC is a probability measure.

*Proof.* First, note that WMC is non-negative and  $\text{WMC}(0) = 0$  by definition. Next, let  $x, y \in \mathbf{B}$  be such that  $x \wedge y = 0$ . We want to show that

$$\text{WMC}(x \vee y) = \text{WMC}(x) + \text{WMC}(y). \quad (1)$$

If, say,  $x = 0$ , then Eq. (1) becomes

$$\text{WMC}(y) = \text{WMC}(0) + \text{WMC}(y) = \text{WMC}(y)$$

(and likewise for  $y = 0$ ). Thus we can assume that  $x \neq 0 \neq y$  and use Lemma 4 to write

$$x = \bigvee_{i \in I} x_i \quad \text{and} \quad y = \bigvee_{j \in J} y_j$$

for some sequences of atoms  $(x_i)_{i \in I}$  and  $(y_j)_{j \in J}$ . If  $x_{i'} = y_{j'}$  for some  $i' \in I$  and  $j' \in J$ , then

$$x \wedge y = \bigvee_{i \in I} \bigvee_{j \in J} x_i \wedge y_j = x_{i'} \wedge y_{j'} \neq 0,$$

contradicting the assumption. This is enough to show that

$$\begin{aligned} \text{WMC}(x \vee y) &= \text{WMC} \left( \left( \bigvee_{i \in I} x_i \right) \vee \left( \bigvee_{j \in J} y_j \right) \right) = \sum_{i \in I} \text{WMC}(x_i) + \sum_{j \in J} \text{WMC}(y_j) \\ &= \text{WMC}(x) + \text{WMC}(y), \end{aligned}$$

finishing the proof that WMC is a measure. This immediately shows that NWMC is a probability measure since, by definition,  $\text{NWMC}(1) = 1$ .  $\square$

Given a theory  $\Delta$  in a logic  $\mathcal{L}$ , the usual way of using WMC to compute the probability of a query  $q$  is [1, 15]

$$\Pr_{\Delta, w}(q) = \frac{\text{WMC}_w(\Delta \wedge q)}{\text{WMC}_w(\Delta)}.$$

In our algebraic formulation, this can be computed in two different ways:

- as  $\frac{\text{WMC}_w(\Delta \wedge q)}{\text{WMC}_w(\Delta)}$  in  $B(\mathcal{L})$ ,
- and as  $\text{NWC}_w([q])$  in  $B(\Delta)$ .

But how does the measure defined on  $B(\mathcal{L})$  transfer to  $B(\Delta)$ ?

## 4 What Measures Are WMC-Definable?

### 4.1 WMC Requires Independent Literals

**Lemma 17.** *For any measure  $m: \mathbf{B} \rightarrow \mathbb{R}_{\geq 0}$  and elements  $a, b \in \mathbf{B}$ ,*

$$m(a \wedge b) = m(a)m(b) \tag{2}$$

*if and only if*

$$m(a \wedge b) \cdot m(\neg a \wedge \neg b) = m(a \wedge \neg b) \cdot m(\neg a \wedge b). \tag{3}$$

*Proof.* First, note that  $a = (a \wedge b) \vee (a \wedge \neg b)$  and  $(a \wedge b) \wedge (a \wedge \neg b) = 0$ , so, by properties of a measure,

$$m(a) = m(a \wedge b) + m(a \wedge \neg b). \tag{4}$$

Applying Eq. (4) and the equivalent expression for  $m(b)$  allows us to rewrite Eq. (2) as

$$m(a \wedge b) = [m(a \wedge b) + m(a \wedge \neg b)][m(a \wedge b) + m(\neg a \wedge b)]$$

which is equivalent to

$$m(a \wedge b)[1 - m(a \wedge b) - m(a \wedge \neg b) - m(\neg a \wedge b)] = m(a \wedge \neg b)m(\neg a \wedge b). \tag{5}$$

Since  $a \wedge b$ ,  $a \wedge \neg b$ ,  $\neg a \wedge b$ ,  $\neg a \wedge \neg b$  are pairwise disjoint and their supremum is 1,

$$m(a \wedge b) + m(a \wedge \neg b) + m(\neg a \wedge b) + m(\neg a \wedge \neg b) = 1,$$

and this allows us to rewrite Eq. (5) into Eq. (3). As all transformations are invertible, the two expressions are equivalent.  $\square$

**Theorem 18.** *Let  $\mathbf{B}$  be a free BA over  $\{l_i\}_{i=1}^n$  (for some  $n \in \mathbb{N}$ ) with measure  $m: \mathbf{B} \rightarrow \mathbb{R}_{\geq 0}$ , and let  $L = \{l_i\}_{i=1}^n \cup \{\neg l_i\}_{i=1}^n$ . Then there exists a weight function  $w: L \rightarrow \mathbb{R}_{\geq 0}$  such that  $m = \text{WMC}_w$  if and only if*

$$m(l \wedge l') = m(l)m(l') \tag{6}$$

*for all distinct  $l, l' \in L$  such that  $l \neq \neg l'$ .*

*Remark.* Note that if  $n = 1$ , then Eq. (6) is vacuously satisfied and so any valid measure can be expressed as WMC.

*Proof.* ( $\Leftarrow$ ) Let  $w: L \rightarrow \mathbb{R}_{\geq 0}$  be defined by

$$w(l) = m(l) \quad (7)$$

for all  $l \in L$ . We are going to show that  $\text{WMC}_w = m$ . First, note that  $\text{WMC}_w(0) = 0 = m(0)$  by the definitions of both  $\text{WMC}_w$  and  $m$ . Second, let

$$a = \bigwedge_{i=1}^n a_i \quad (8)$$

be an atom in  $\mathbf{B}$  such that  $a_i \in \{l_i, \neg l_i\}$  for all  $i \in [n]$ . Then

$$\text{WMC}(a) = \prod_{i=1}^n w(a_i) = \prod_{i=1}^n m(a_i) = m\left(\bigwedge_{i=1}^n a_i\right) = m(a)$$

by Definition 15 and Eqs. (6) to (8). Finally, note that if  $\text{WMC}$  and  $m$  agree on all atoms, then they must also agree on all other non-zero elements of the Boolean algebra.

( $\Rightarrow$ ) For the other direction, we are given a weight function  $w: L \rightarrow \mathbb{R}_{\geq 0}$  that induces a measure  $m = \text{WMC}_w: \mathbf{B} \rightarrow \mathbb{R}_{\geq 0}$ , and we want to show that Eq. (6) is satisfied. Let  $k_i, k_j \in L$  be such that  $k_i \in \{l_i, \neg l_i\}$ ,  $k_j \in \{l_j, \neg l_j\}$ , and  $i \neq j$  for some  $i, j \in [n]$ . We then want to show that

$$m(k_i \wedge k_j) = m(k_i)m(k_j) \quad (9)$$

which is equivalent to

$$m(k_i \wedge k_j) \cdot m(\neg k_i \wedge \neg k_j) = m(k_i \wedge \neg k_j) \cdot m(\neg k_i \wedge k_j) \quad (10)$$

by Lemma 17. Then

$$\begin{aligned} \text{WMC}(k_i \wedge k_j) &= \sum_{\text{atoms } a \leq k_i \wedge k_j} \text{WMC}(a) = \sum_{\text{atoms } a \leq k_i \wedge k_j} \prod_{m \in [n]} w(a_m) \\ &= \sum_{\text{atoms } a \leq k_i \wedge k_j} w(a_i)w(a_j) \prod_{m \in [n] \setminus \{i, j\}} w(a_m) = \sum_{\text{atoms } a \leq k_i \wedge k_j} w(k_i)w(k_j) \prod_{m \in [n] \setminus \{i, j\}} w(a_m) \\ &= w(k_i)w(k_j) \sum_{\text{atoms } a \leq k_i \wedge k_j} \prod_{m \in [n] \setminus \{i, j\}} w(a_m) = w(k_i)w(k_j)C, \end{aligned}$$

where  $C$  denotes the part of  $\text{WMC}(k_i \wedge k_j)$  that will be the same for  $\text{WMC}(\neg k_i \wedge k_j)$ ,  $\text{WMC}(k_i \wedge \neg k_j)$ , and  $\text{WMC}(\neg k_i \wedge \neg k_j)$  as well. But then Eq. (10) becomes

$$w(k_i)w(k_j)w(\neg k_i)w(\neg k_j)C^2 = w(k_i)w(\neg k_j)w(\neg k_i)w(k_j)C^2$$

which is trivially true.  $\square$

## 4.2 Extending the Algebra

Given this requirement for independence, a well-known way to represent probability distributions that do not consist entirely of independent variables is by adding more literals [3], i.e., extending the set  $L$  covered by the  $\text{WMC}$  weight function  $w: L \rightarrow \mathbb{R}_{\geq 0}$ . Let us translate this idea to the language of BAs.

**Theorem 19.** *Let  $\mathbf{B}$  be a free BA over a finite set  $S$ , and let  $m: \mathbf{B} \rightarrow \mathbb{R}_{\geq 0}$  be an arbitrary measure. Let  $L = \{s \mid s \in S\} \cup \{\neg s \mid s \in S\}$ . By Lemma 12, we know that  $\mathbf{B}$  has  $n = 2^{|S|}$  atoms. Let  $\{a_i\}_{i=1}^n$  denote those atoms in some arbitrary order. Let  $L' = L \cup \{\phi_i\}_{i=1}^n \cup \{\neg \phi_i\}_{i=1}^n$  be the set  $L$  extended with  $2n$  new elements. Let  $\mathbf{B}'$  be the unique Boolean algebra with  $\{\phi_i \wedge a_i\}_{i=1}^n \cup \{\neg \phi_i \wedge a_i\}_{i=1}^n$  as its set of atoms. Let  $\iota: \mathbf{B} \hookrightarrow \mathbf{B}'$  be the inclusion homomorphism. Let  $w: L' \rightarrow \mathbb{R}_{\geq 0}$  be defined by*

$$w(l) = \begin{cases} m(a_i)/2 & \text{if } l = \phi_i \text{ or } l = \neg \phi_i \text{ for some } i \in [n] \\ 1 & \text{otherwise} \end{cases}$$

for all  $l \in L'$ , and note that this defines a WMC measure  $m' = \text{WMC}_w : \mathbf{B}' \rightarrow \mathbb{R}_{\geq 0}$ . Then

$$m(a) = (m' \circ \iota)(a)$$

for all  $a \in \mathbf{B}$ .

In other words, any measure can be computed using WMC by extending the BA with more literals. More precisely, we are given the left-hand column in

$$\begin{array}{ccc} \mathbb{R}_{\geq 0} & & \\ \uparrow m & \swarrow m' & \\ \mathbf{B} & \xrightarrow{\iota} & \mathbf{B}' \\ \uparrow & & \uparrow \\ L & \hookrightarrow & L' \xrightarrow{w} \mathbb{R}_{\geq 0} \end{array}$$

and construct the remaining part in such a way that the triangle commutes.

*Proof.* Since  $\mathbf{B}$  is free over  $S$ , each atom  $a_i \in \mathbf{B}$  is an infimum of elements in  $L$ , i.e.,

$$a_i = \bigwedge_{j \in J} a_{i,j}$$

for some  $\{a_{i,j}\}_{j \in J} \subset L$ . Moreover, each atom  $b \in \mathbf{B}'$  can be represented as either  $b = \phi_i \wedge a_i$  or  $b = \neg\phi_i \wedge a_i$  for some atom  $a_i \in \mathbf{B}$ , also making it an infimum over a subset of  $L'$ . Then, for any  $b \in \mathbf{B}$ ,

$$(m' \circ \iota)(b) = \sum_{\substack{\text{atoms } a_i \in \mathbf{B}: \\ \phi_i \wedge a_i \leq \iota(b)}} (w(\phi_i) + w(\neg\phi_i)) \prod_{j \in J} w(a_{i,j}),$$

recognising that, for any  $\iota(b)$ , any atom  $a_i \in \mathbf{B}$  satisfies  $\phi_i \wedge a_i \leq \iota(b)$  if and only if it satisfies  $\neg\phi_i \wedge a_i \leq \iota(b)$ . Then, according to the definition of  $w$ ,

$$(m' \circ \iota)(b) = \sum_{\substack{\text{atoms } a_i \in \mathbf{B}: \\ \phi_i \wedge a_i \leq \iota(b)}} (w(\phi_i) + w(\neg\phi_i)) = \sum_{\substack{\text{atoms } a_i \in \mathbf{B}: \\ \phi_i \wedge a_i \leq \iota(b)}} m(a_i) = m(b),$$

provided that

$$\phi_i \wedge a_i \leq \iota(b) \quad \text{if and only if} \quad a_i \leq b,$$

but this is equivalent to

$$\phi_i \wedge a_i = \phi_i \wedge a_i \wedge b \quad \text{if and only if} \quad a_i = a_i \wedge b$$

which is true because  $\phi_i \notin L$ . □

Now we can show that the construction in Theorem 19 is smallest possible.

**Conjecture 20.** Let  $\mathbf{B}$  and  $\mathbf{B}'$  be Boolean algebras, and  $\iota : \mathbf{B} \hookrightarrow \mathbf{B}'$  be the inclusion map such that  $\mathbf{B}$  is free over  $L$ , all atoms of  $\mathbf{B}'$  can be expressed as meets of elements of  $L'$ , and the following subset relations are satisfied:

$$\begin{array}{ccc} \mathbf{B} & \xhookrightarrow{\iota} & \mathbf{B}' \\ \cup & & \cup \\ L & \subset & L' \end{array}$$

If, for any measure  $m : \mathbf{B} \rightarrow \mathbb{R}_{\geq 0}$ , one can construct a weight function  $w : L' \rightarrow \mathbb{R}_{\geq 0}$  such that the WMC measure  $\text{WMC} : \mathbf{B}' \rightarrow \mathbb{R}_{\geq 0}$  with respect to  $w$  satisfies

$$m = \text{WMC} \circ \iota,$$

then  $|L' \setminus L| \geq 2^{|L|+1}$ .

Let us note how our lower bound on the number of added literals compares to two methods of translating a discrete probability distribution into a WMC problem over a propositional knowledge base proposed by Darwiche [4] and Sang et al. [15]. Suppose we have a discrete probability distribution with  $n$  variables, and the  $i$ th variable has  $v_i$  values, for each  $i \in [n]$ . Interpreted as a logical system, it has  $\prod_{i=1}^n v_i$  models. My expansion would then use

$$\sum_{i=1}^n v_i + 2 \prod_{i=1}^n v_i$$

variables, i.e., a variable for each possible variable-value assignment, and two additional variables for each model. Without making any independence assumptions, the encoding by Darwiche [4] would use

$$\sum_{i=1}^n v_i + \sum_{i=1}^n \prod_{j=1}^i v_j$$

variables, while for the encoding by Sang et al. [15],

$$\sum_{i=1}^n v_i + \sum_{i=1}^n (v_i - 1) \prod_{j=1}^{i-1} v_j$$

variables would suffice.

## 5 Representing Independence and Conditional Independence

### 5.1 Preliminaries

**Definition 21.** Given a BA  $\mathbf{A}$ , a *subalgebra* is a subset  $\mathbf{B} \subseteq \mathbf{A}$  that, together with the operations, zero, and one of  $\mathbf{A}$ , is a BA.

**Definition 22** ([8]). Let  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{C}$  be BAs such that  $\mathbf{B}$  is a subalgebra of  $\mathbf{A}$ . Let  $f: \mathbf{A} \rightarrow \mathbf{C}$  and  $g: \mathbf{B} \rightarrow \mathbf{C}$  be homomorphisms. Then  $f$  is an *extension* of  $g$  if  $f(x) = g(x)$  for all  $x \in \mathbf{B}$ . If  $f$  is an extension of each member of a family  $\{g_i\}_{i \in I}$  of homomorphisms, then  $f$  is called a *common extension* of  $\{g_i\}_{i \in I}$ .

**Definition 23** ([8]). Let  $\{\mathbf{A}_i\}_{i \in I}$  be a family of subalgebras of a BA  $\mathbf{A}$ . If for any BA  $\mathbf{B}$  with a family of homomorphisms  $\{f_i: \mathbf{A}_i \rightarrow \mathbf{B}\}_{i \in I}$  there exists a unique common extension of  $\{f_i: \mathbf{A}_i \rightarrow \mathbf{B}\}_{i \in I}$  ( $f: \mathbf{A} \rightarrow \mathbf{B}$  in the diagram),

$$\begin{array}{ccc} \mathbf{A}_i & \hookrightarrow & \mathbf{A} \\ & \searrow f_i & \downarrow f \\ & & \mathbf{B} \end{array}$$

then  $\mathbf{A}$  is the *internal sum*<sup>3</sup> of  $\{\mathbf{A}_i\}_{i \in I}$ . We will denote it as  $\bigoplus_{i \in I} \mathbf{A}_i$ .

**Proposition 24** ([16]). Let  $\mathbf{A}$  be the internal sum of a family of BAs  $\{\mathbf{A}_i\}_{i \in I}$ , and let  $\{m_i: \mathbf{A}_i \rightarrow \mathbb{R}_{\geq 0}\}_{i \in I}$  be a family of measures. Then there exists a unique measure  $m: \mathbf{A} \rightarrow \mathbb{R}_{\geq 0}$  such that, for any finite subset  $J \subseteq I$  and family of elements  $\{x_j \in \mathbf{A}_j\}_{j \in J}$ ,

$$m \left( \bigwedge_{j \in J} x_j \right) = \prod_{j \in J} m_j(x_j).$$

<sup>3</sup>A slightly more general version of this definition is also known as the free product, the Boolean product, and the coproduct in the category of BAs [8, 11, 16].



**Definition 25** ([11]). Let  $\mathbf{A}$  be a BA. Let  $\mathbf{B}$  be a subalgebra of  $\mathbf{A}$ , and let  $\{\mathbf{A}_i\}_{i \in I}$  be a family of subalgebras of  $\mathbf{A}$  such that  $\mathbf{A}_i \cap \mathbf{A}_j = \mathbf{B}$  for all  $i \neq j$  in  $I$ . Let  $\{\iota_i: \mathbf{B} \rightarrow \mathbf{A}_i\}$  be a family of inclusion homomorphisms. Then  $\mathbf{A}$  is the *amalgamated free product*<sup>4</sup> of  $\{\mathbf{A}_i\}_{i \in I}$  over  $\mathbf{B}$  if, for any Boolean algebra  $\mathbf{C}$  with a family of homomorphisms  $\{f_i: \mathbf{A}_i \rightarrow \mathbf{C}\}_{i \in I}$  such that  $f_i \circ \iota_i = f_j \circ \iota_j$  for all  $i, j \in I$ , there is a unique homomorphism  $f: \mathbf{A} \rightarrow \mathbf{C}$  such that the triangle in

$$\begin{array}{ccccc} \mathbf{B} & \xrightarrow{\iota_i} & \mathbf{A}_i & \hookrightarrow & \mathbf{A} \\ & & \downarrow f_i & \swarrow f & \\ & & \mathbf{C} & & \end{array}$$

commutes for all  $i \in I$ . We will denote this product as

$$\mathbf{A} = \bigoplus_{\substack{\mathbf{B} \\ i \in I}} \mathbf{A}_i.$$

## 5.2 New Results

**Theorem 26.** Let  $\{S_i\}_{i=0}^n$  be a finite set of finite sets for some  $n > 1$  such that for all distinct positive integers  $i$  and  $j$ ,  $S_i \cap S_j = S_0$ , and let

$$\mathbf{A} = \bigoplus_{\substack{\mathcal{F}(S_0) \\ 1 \leq i \leq n}} \mathcal{F}(S_i).$$

Let  $(m_i: \mathcal{F}(S_i) \rightarrow \mathbb{R}_{\geq 0})_{i=1}^n$  be arbitrary measures. Then there is a unique measure  $m: \mathbf{A} \rightarrow \mathbb{R}_{\geq 0}$  such that, for any element  $b \in \mathcal{F}(S_0)$ , subset  $J \subseteq \{1, 2, \dots, n\}$ , and elements  $\{a_j \in \mathcal{F}(S_j \setminus S_0)\}_{j \in J}$ ,

$$m \left( b \wedge \bigwedge_{j \in J} a_j \right) = \prod_{j \in J} m_j(b \wedge a_j).$$

*Proof.* □

**Theorem 27.** The number of weights needed to encode a Bayesian network using coproducts and pushouts is equal to the number of entries in the tables of the network (and the resulting theory is shorter).

**Theorem 28** (Pushouts of free BAs are free). Let

$$\mathbf{A} = \bigoplus_{\substack{\mathbf{B} \\ i \in I}} \mathbf{A}_i$$

be an amalgamated free product such that  $\{\mathbf{A}_i\}_{i \in I}$  are free BAs with  $\{S_i\}_{i \in I}$  as their respective sets of generators. Let  $S = \bigcup_{i \in I} S_i$ . Then  $\mathbf{A}$  is a free BA with generating set  $S$ .

*Proof.* Suppose we have a map from  $S$  to an arbitrary BA  $\mathbf{C}$ , as in

$$\begin{array}{ccc} S_i & \hookrightarrow & S \\ \downarrow & & \downarrow \\ \mathbf{A}_i & \hookrightarrow & \mathbf{A} \end{array} \quad \begin{array}{c} \searrow \\ \downarrow \\ \mathbf{C} \end{array}$$

We want to show that there exists a unique homomorphism  $\mathbf{A} \rightarrow \mathbf{C}$ . For all  $i \in I$ , from  $S_i \hookrightarrow S$  and  $S \rightarrow \mathbf{C}$  we get a map  $S_i \rightarrow \mathbf{C}$ , so—by the definition of a free BA—there is a unique homomorphism  $\mathbf{A}_i \rightarrow \mathbf{C}$ . Furthermore, a family of homomorphisms  $\{\mathbf{A}_i \rightarrow \mathbf{C}\}_{i \in I}$  uniquely determine a homomorphism  $\mathbf{A} \rightarrow \mathbf{C}$  by the universal mapping property of a (wide) pushout. Thus  $\mathbf{A}$  is a free BA with generating set  $S$ . □

<sup>4</sup>Also known as a (wide) pushout in the category of BAs.

**Corollary 29.** *Similarly, coproducts of free BAs are free.*

## 6 How to represent a Bayesian network

Let  $G = (V, E)$  be a DAG, and let  $(X_v)_{v \in V}$  be the discrete random variables (RVs). Let  $(d_v)_{v \in V}$  denote the numbers of values that each RV can take and assume that  $d_v$  is a power of two for all  $v \in V$ .

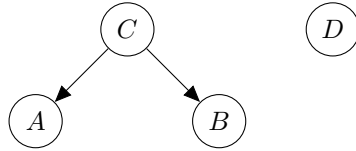
1. For each  $v \in V$ , let  $A_v$  be a set of  $\log_2 d_v$  Boolean variables. (If  $d_v$  is not a power of two, we would need to manually forbid each illegal value, e.g., by adding formulas such as  $\neg(x_1 \wedge x_2 \wedge x_3 \wedge \neg x_4)$  to the theory. Alternatively, we could set their probability to zero, but the former should be better.)
2. For  $v \in V$ :
  - (a) If  $v$  has no parents, we can directly provide a probability measure on  $\mathcal{F}(A_v)$ .
  - (b) Otherwise, we need a probability measure on

$$\mathcal{F}\left(A_v \cup \bigcup_{(u,v) \in E} A_u\right).$$

For any element  $a_u \wedge a_v$  in this free Boolean algebra such that  $a_v \in A_v \setminus \{0, 1\}$ , and  $a_u \notin A_v$ , we can set

$$m(a_u \wedge a_v) = \Pr(a_v \mid a_u) \Pr(a_u).$$

3. When a #SAT algorithm asks for the weight of a model  $\bigwedge_{v \in V} l_v$ , split it into segments that correspond to each leaf, take the probabilities from each such node, and multiply them. NB: The segments will intersect, e.g., with



$a \wedge b \wedge c \wedge d$  will be split into  $(a \wedge c) \wedge (b \wedge c) \wedge d$ .

### 6.1 Bayesian networks as Boolean algebras




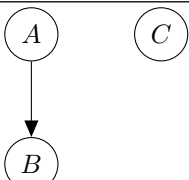
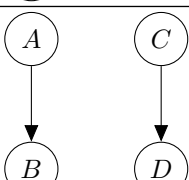
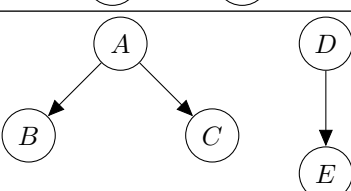
To turn a Bayesian network into a Boolean algebra, run  $f(\{v \in V \mid \text{outdeg}(v) = 0\})$ , where  $f$  works as follows on an arbitrary  $A \subseteq V$ :

1.  $S \leftarrow \{u \mid (u, v) \in E, v \in A\}$  (i.e.,  $S$  is the set of parents of the elements in  $A$ ).
2. If  $S = \emptyset$ , return  $\sum A$  (a coproduct).
3. Otherwise, return  $\sum_{f(S)} A$  (a pushout).

## References

- [1] Vaishak Belle. Weighted model counting with function symbols. In Gal Elidan, Kristian Kersting, and Alexander T. Ihler, editors, *Proceedings of the Thirty-Third Conference on Uncertainty in Artificial Intelligence, UAI 2017, Sydney, Australia, August 11-15, 2017*. AUAI Press, 2017.
- [2] Elena Castiñeira, Susana Cubillo, and Enric Trillas. On possibility and probability measures in finite Boolean algebras. *Soft Comput.*, 7(2):89–96, 2002.

Table 1: Examples of how Bayesian networks encode as Boolean algebras. Here,  $A$  denotes  $\mathcal{F}\{a\}$  and  $AB$  denotes  $\mathcal{F}\{a, b\}$ .

Bayesian network	Boolean algebra
	$A$
	$A + B$
	$AB$
	$B +_A C$
	$B +_{A+C} D$
	$B +_{A+D} C +_{A+D} E$

- [3] Mark Chavira and Adnan Darwiche. On probabilistic inference by weighted model counting. *Artif. Intell.*, 172(6-7):772–799, 2008.
- [4] Adnan Darwiche. A logical approach to factoring belief networks. In Dieter Fensel, Fausto Giunchiglia, Deborah L. McGuinness, and Mary-Anne Williams, editors, *Proceedings of the Eight International Conference on Principles and Knowledge Representation and Reasoning (KR-02), Toulouse, France, April 22-25, 2002*, pages 409–420. Morgan Kaufmann, 2002.
- [5] Haim Gaifman. Concerning measures on Boolean algebras. *Pacific Journal of Mathematics*, 14(1):61–73, 1964.
- [6] M. Ganesh. *Introduction to fuzzy sets and fuzzy logic*. PHI Learning Pvt. Ltd., 2006.
- [7] Scott Garrabrant, Tsvi Benson-Tilsen, Andrew Critch, Nate Soares, and Jessica Taylor. Logical induction. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:154, 2016.
- [8] Steven Givant and Paul R. Halmos. *Introduction to Boolean algebras*. Springer Science & Business Media, 2008.
- [9] Theodore Hailperin. Probability logic. *Notre Dame Journal of Formal Logic*, 25(3):198–212, 1984.
- [10] Thomas Jech. *Set theory, Second Edition*. Perspectives in Mathematical Logic. Springer, 1997.
- [11] Sabine Koppelberg, Robert Bonnet, and James Donald Monk. *Handbook of Boolean algebras*, volume 384. North-Holland Amsterdam, 1989.
- [12] Peter H. Krauss. Representation of conditional probability measures on Boolean algebras. *Acta Mathematica Hungarica*, 19(3-4):229–241, 1968.
- [13] Ken Levasseur and Al Doerr. *Applied Discrete Structures*. Lulu.com, 2012.
- [14] Nils J. Nilsson. Probabilistic logic. *Artif. Intell.*, 28(1):71–87, 1986.
- [15] Tian Sang, Paul Beame, and Henry A. Kautz. Performing Bayesian inference by weighted model counting. In Manuela M. Veloso and Subbarao Kambhampati, editors, *Proceedings, The Twentieth National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference, July 9-13, 2005, Pittsburgh, Pennsylvania, USA*, pages 475–482. AAAI Press / The MIT Press, 2005.
- [16] Roman Sikorski. *Boolean algebras*. Springer, third edition, 1969.
- [17] Alfred Tarski. *Logic, semantics, metamathematics: papers from 1923 to 1938*. Hackett Publishing, 1983.